Arabian Journal of Chemistry (2015) xxx, xxx–xxx

King Saud University

Arabian Journal of Chemistry

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Molecular modeling study of uracil-based hydroxamic acids-containing histone deacetylase inhibitors

Mukesh C. Sharma [a], **Smita Sharma** [b,*]

[a] *School of Pharmacy, Devi Ahilya University, Takshila Campus, Khandwa Road, Indore, M.P. 452 001, India*
[b] *Department of Chemistry, Chodhary Dilip Singh Kanya Mahavidyalaya, Bhind, M.P. 477 001, India*

Received 25 July 2013; accepted 31 December 2014

KEYWORDS

Uracil;
Histone deacetylase;
Maize deacetylases (HD2);
2D-QSAR;
Molecular field analysis;
Partial least squares

Abstract A quantitative structure activity relationship study was performed on a series of uracil based hydroxamide Inhibitors of maize deacetylases HD2 and histone deacetylase (Mouse HDAC1) activity for establishing quantitative relationship between biological activity and their physicochemical properties. The two dimensional and *k-nearest neighbor* studies were performed using partial least square methodology coupled with genetic algorithm (GA) and simulated annealing (SA) was applied to derive models. The 2D model developed gave good correlation coefficient (r^2) of 0.8498, and r^2 for external test set ($pred\_r^2$) 0.7932 was developed by GA-PLS with the descriptors such as Hydrogen count, SsssCHcount, and SdsCHE-index. *k-nearest neighbor* method was applied for the generation of steric and electrostatic descriptors based on aligned structures. 3D QSAR studies produced reasonably good predictive models with high cross-validated q^2 value of 0.679 and $pred\_r^2$ = 0.733 values using the model GA kNN-MFA method. The best pharmacophore shows that the four features used were one AroC feature (Aromatic), one AlaC (aliphatic) and two HAc (Hydrogen bond acceptor) features. The average RMSD of the pharmacophore alignment of each two molecules is 0.0754 Å. The QSAR result gives relationship between structural Uracil-based hydroxamic acids derivatives and their activities which should be useful to design newer histone deacetylase inhibitors.

1. Introduction

Histone deacetylases (HDACs) represent a family of enzymes that compete with histone acetyltransferases (HATs) to modulate chromatin structure and transcriptional activity via change in acetylation status of nucleosomal histones. HDACs are deacetylating the ε-amino groups of lysine located near the amino termini of core histone proteins (Monneret, 2005;

* Corresponding author.
E-mail addresses: mukeshcsharma@yahoo.com (M.C. Sharma), drsmita.sharma@rediffmail.com (S. Sharma).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.arabjc.2014.12.030

Mai et al., 2002). To date, four classes of mammalian HDACs are known according to their homology with the corresponding yeast transcriptional regulators. Class I (HDAC1-3,8), IIa (HDAC4,5,7,9), IIb (HDAC6,-10), and IV (HDAC11) HDACs are Zn^{2+}-dependent deacetylases, are components of multiprotein complexes containing other proteins known to function in transcriptional activation/repression, and differ for their subcellular localization and tissue expression (Grozinger and Schreiber, 2002; Gregoretti et al., 2004; Verdin et al., 2003). Acetylation and deacetylation of the specific lysines within histones play a crucial role in the transcriptional process (Pazin and Kadonaga, 1997). Two families of enzymes, acetylases and deacetylases, are involved in controlling the acetylation state of histones. Histone deacetylase (HDAC) enzymes, which regulate the level of histone acetylation, are one of the major groups mediating epigenetic control. HDAC enzymes act by regulating the level of biological acetylation and deacetylation reactions of their targets. Acetylation of the N-terminal region of histone proteins promotes gene expression (Garea and Esteller, 2004; Somech et al., 2004). Recent studies show that inhibition of histone deacetylases elicits anticancer effects in several tumor cells by inhibition of cell growth and induction of cell differentiation. The development of HDAC inhibitors as anticancer drugs has been initiated, and compounds such as the hydroxamic acid Trichostatin A (TSA) (Yoshida et al., 1990) suberanilohydroxamic acid (Richon et al., 1998) the cyclic tetrapeptides apicidin (Han et al., 2000) and trapoxin, (Kijima et al., 1993) as well as synthetic inhibitors have been studied for this purpose in cancer cell lines. Thus, inhibition of HDACs, which induces histone hyperacetylation, provides a potential target for the development of synthetic anticancer drugs (Bouchain et al., 2001; Curtin and Glaser, 2003; Weinmann and Ottow, 2004). Histone deacetylase inhibitors (HDACis) exert cell type – specific effects including apoptosis, cell-cycle arrest and differentiation. In leukemias, HDACis include the expression of members of the tumor-necrosis factor-related apoptosis-inducing ligand and FAS death receptor pathways. This induction is responsible for the pro-apoptotic efforts of HDACis (Johnstone, 2002; Marks et al., 2001; Insinga et al., 2005; Nebbioso et al., 2005). Cell-based studies have shown that HDACIs have a powerful antiproliferative property, causing cell-cycle arrest, apoptosis, and differentiation; these antiproliferative effects are far more pronounced in tumor cells than in normal cells (Inche and La Thangue, 2006). HDACs have increasingly become important targets and the hunt for HDAC inhibitors has been intensified and attracted great attention in drug discovery over the years (Pandolfi, 2001). 2D-QSAR relationship is a rough approximation and contains topological or two-dimensional (2D) information. It explains how the atoms are bonded in a molecule, the type of bonding, and the interaction of particular atoms (e.g., total path count, molecular connectivity indices, etc.). The pharmacophore modeling is a well established approach to quantitatively explore common chemical features among a considerable number of structures and qualified pharmacophore model could also be used as a query for searching chemical databases to find new chemical entities (Sotriffer et al., 1996).

The present work is an attempt to generate predictive QSAR models based on 2D and 3D-QSAR methods and to find the structural features of uracil-based hydroxamic acids required for histone deacetylase inhibitor activities to guide the rational synthesis of novel histone deacetylase inhibitors. In this investigation, widely used technique, viz., genetic algorithm (GA) and simulated annealing (SA) has been applied for descriptor optimization, and kNN-MFA analysis has been applied for 3D-QSAR model development.

2. Materials and methods

2.1. Dataset for analysis

A data set of thirty-four N-hydroxy-polymethylenealkanamide group (uracil-based hydroxamic acids) derivatives of reported series for histone deacetylase inhibitors was used for the present QSAR study (Mai et al., 2006). The biological activity values [IC_{50} (nM)] reported in the literature were converted to their molar units and then further to negative logarithmic scale (pIC_{50}) and subsequently used as the dependent variable for the QSAR analysis. Table 1 shows the structure of 34 such compounds along with their biological activity values.

2.2. Selection of training and test set

The total set of 34 compounds was divided into a training set (27 compounds) for generating 2D QSAR models and a test set (7 compounds) for validating the quality of the models (Table 1). Selection of the training set and test set molecules was done on the basis of structural diversity and a wide range of activity such that the test-set molecules represent a range of biological activity similar to that of the training set; thus, the test set is truly representative of the training set (Golbraikh and Tropsha, 2002). The unicolumn statistics of the training and test sets is reported in Table 2. The maximum and minimum values in training and test set were compared in a way that:

1. The maximum value of pIC_{50} of test set should be less than or equal to maximum value of pIC_{50} of training set.
2. The minimum value of pIC_{50} of test set should be higher than or equal to minimum value of pIC_{50} of training set.

This observation showed that test set was interpolative and derived within the minimum–maximum range of training set. The mean and standard deviation of pIC_{50} values of sets of training and test provide insights into the relative difference of mean and point density distribution (along mean) of the two sets (VLife MDS, 2008).

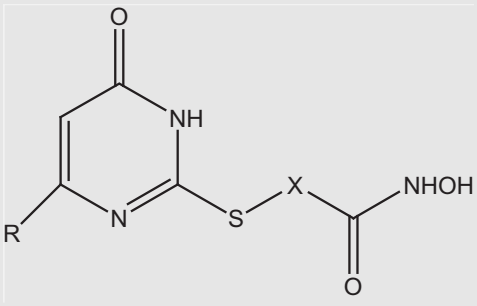
2.3. Calculation of 2D-QSAR descriptors

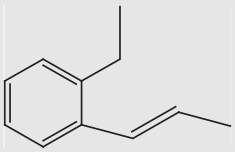
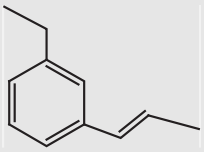
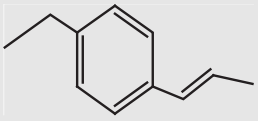
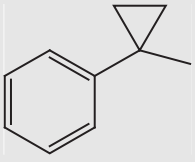
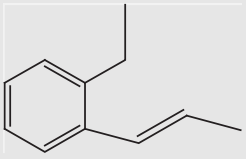
The molecular structures of all the 34 molecules were built using the 2D draw application of V-Life MDS 3.5 software (VLife MDS, 2008) with standard bond lengths and bond angles. Geometry optimization was carried out using the standard Merck Molecular Force Field (MMFF) followed by considering distance-dependent dielectric constant of 1.0, convergence criterion or root-mean-square (RMS) gradient at 0.01 kcal/mol Å and the iteration limit to 10,000 (Halgren, 1996). The energy-minimized geometry was used for the calculation of the molecular descriptors.

A total of 264 descriptors were calculated by using VLife Sciences Molecular Design Suite which was subsequently

Please cite this article in press as: Sharma, M.C., Sharma, S. Molecular modeling study of uracil-based hydroxamic acids-containing histone deacetylase inhibitors. Arabian Journal of Chemistry (2015), http://dx.doi.org/10.1016/j.arabjc.2014.12.030

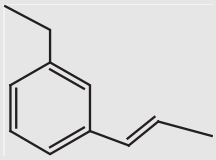
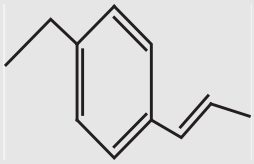
Table 1 Structure, and biological activity of uracil-containing histone deacetylase inhibitors.



Com. no.	R	X	IC_{50}	pIC_{50}
1	Ph	$(CH_2)_2$	822	2.914
2	Ph	$(CH_2)_3$	27	1.431
3[a]	Ph	$(CH_2)_4$	8	0.903
4	Ph	$(CH_2)_5$	12	1.079
5	Ph	$(CH_2)_6$	38	1.579
6	Ph	$(CH_2)_7$	42	1.623
7	Ph		367	2.564
8	Ph		17	1.230
9[a]	Ph		23	1.361
10	Ph CH_2	CH=CH	9000	3.954
11[a]	Ph CH_2	$(CH_2)_2$	38	1.579
12	Ph CH_2	$(CH_2)_3$	229	2.359
13	Ph CH_2	$(CH_2)_4$	125	2.096
14	Ph CH_2	$(CH_2)_5$	18	1.255
15[a]	$PhCH(CH_3)$	$(CH_2)_5$	9	0.954
16	$PhCH(C_2H_5)$	$(CH_2)_5$	32	1.505
17	$PhCH(OCH_3)$	$(CH_2)_5$	82	1.913
18		$(CH_2)_5$	62	1.792
19	PhCH(Ph)	$(CH_2)_5$	52	1.716
20[a]	$PhCH_2$	$(CH_2)_6$	37	1.568
21	$PhCH_2$	$(CH_2)_7$	61	1.785
22	$PhCH_2$		760	2.880

(*continued on next page*)

Table 1 (*continued*)

Com. no.	R	X	IC_{50}	pIC_{50}
23[a]	$PhCH_2$		41	1.612
24	$PhCH_2$		80	1.903
25	$PhCH_2CH_2$	CH═CH	41,000	4.612
26[a]	$PhCH_2CH_2$	$(CH_2)_2$	37	1.568
27	$PhCH_2CH_2$	$(CH_2)_3$	205	2.311
28	$PhCH_2CH_2$	$(CH_2)_4$	40	1.602
29	$PhCH_2CH_2$	$(CH_2)_5$	35	1.544
30	$PhCH_2CH_2$	$(CH_2)_6$	83	1.919
31	$PhCH_2CH_2$	$(CH_2)_7$	90	1.954
32	H	$(CH_2)_5$	213	2.328
33	Me	$(CH_2)_5$	110	2.041
34	n-Pr	$(CH_2)_5$	135	2.130

[a] The compounds considered in the test set in 2D QSAR and 3D QSAR.

Table 2 Unicolumn statistics of the training and test sets for activity.

Data set	Average	Max	Min	Std. dev.	Sum
Training	1.9377	4.6128	0.9031	0.8995	44.563
Test	1.8140	2.5647	1.2304	0.4617	14.5122

reduced to 231 descriptors. The descriptors having the same value or almost same value or highly correlated with other descriptors were removed initially. For calculation of AI descriptors (Baumann, 2002) every atom in the molecule was assigned at least one and at most three attributes.

In this study to calculate AI descriptors, we have used following attributes, 2 (double bonded atom), 3 (triple bonded atom), C, N, O, S, H, F, Cl, Br and I and the distance range of 0–7. A value of $pred\_r^2$ greater than 0.5 indicates the good predictive capacity of the QSAR model. However, a QSAR model is considered to be predictive, if the following conditions are satisfied: $r^2 > 0.6$, $q^2 > 0.6$ and $pred\_r^2 > 0.5$ (Golbraikh and Tropsha, 2002).

2.4. Calculation of 3D-QSAR descriptors

Energy minimized and geometry optimized structure of molecules were aligned by the template-based method (Ajmani et al., 2006). The template structure, i.e., Uracil-based 2-phenylethyl moiety ring was used for alignment by considering the common elements of the series as shown in Fig. 1a. The compound 8 possessed very high Histone deacetylase activity which made it a valid lead molecule and therefore was chosen as a reference molecule. The superimposition of all molecules is shown in Fig. 1b.

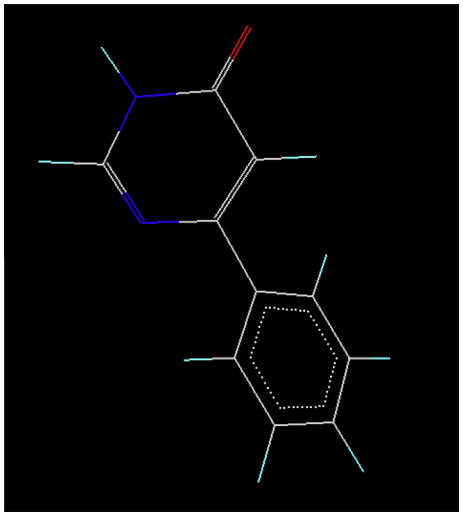


Figure 1a Common template view for aligned molecules.

For calculation of field descriptor values, using *Tripos force field* (Clark et al., 1989), steric, electrostatic and hydrophobic field types, with cutoffs of 10.0 and 30.0 kcal/mol, were selected and charge type was selected as by Gasteiger and Marsili (1980). The *k-nearest neighbor* field depicting the steric, electrostatic and hydrophobic interaction with methyl probe with +1.0 charges was calculated using MMFF. This resulted in

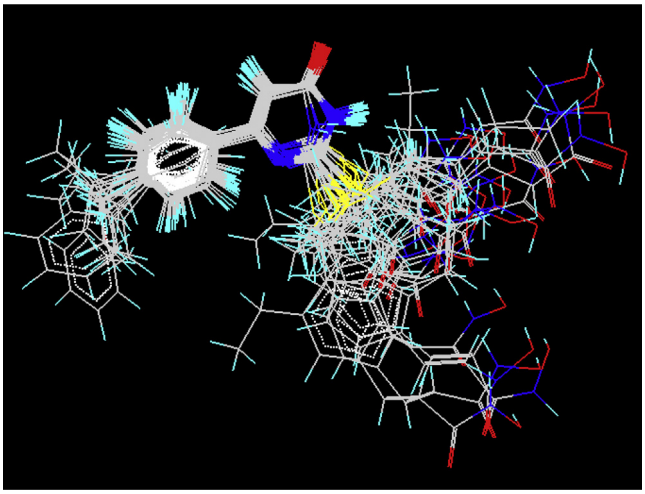


Figure 1b Molecular alignment of the compounds.

calculation of 7500 field descriptors (2500 for each steric, electrostatic and hydrophobic which theoretically form a continuum) for all the compounds in separate columns after removing descriptors having zero values or same values (Table 3).

2.5. Pharmacophore model generation

This study was performed using the software package VLife MDS 3.5. Mol-Sign Module is used for the identification, generation and analysis of pharmacophore by aligning small organic molecules based on their 3D pharmacophore features. For four point pharmacophore identification tolerance limit set up to 30 Å and max distance allowed between two features, set the value to 5 Å.

2.6. Model validation

Internal validation was carried out using leave-one-out (q^2, LOO) method (Cramer et al., 1988). To calculate q^2, each molecule in the training set was sequentially removed, the model refit using same descriptors, and the biological activity of the removed molecule predicted using the refit model. The q^2 was calculated using Eq. (1):

$$q^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{mean})^2} \quad (1)$$

where $y_i, \hat{y}_i$ are the actual and predicted activity of the ith molecule in the training set, respectively, and y_{mean} is the average activity of all molecules in the training set. For external

Table 3 Selected descriptor parameters of uracil-containing histone deacetylase inhibitors.

Hydrogen count	$SsCH_3E$-index	$SssCH_2$count	E_685	S_1300	S_723
13	1	0	−4.06345	−0.0815	−0.03115
17	1	3	−2.55971	−0.06849	−0.02768
19	2	2	−3.11461	−0.08471	−0.03474
13	1	1	−3.15068	−0.08472	−0.03406
21	3	1	−3.57275	−0.0712	−0.02986
23	2	4	−5.44394	−0.09349	−0.04915
25	3	3	−4.19008	−0.07958	−0.03186
23	3	2	−2.59508	−0.08244	−0.03261
23	0	5	−2.60821	−0.0899	−0.03615
23	2	1	−2.56547	−0.07972	−0.0327
15	0	3	−4.12765	−0.08894	−0.0361
23	3	2	−3.71669	−0.08671	−0.0366
23	1	6	−3.22463	−0.085	−0.03714
25	2	3	−2.4758	−0.06276	−0.02702
25	2	3	−3.33551	−0.06843	−0.02764
25	2	3	−3.18016	−0.08759	−0.03677
17	1	2	−3.66436	−0.08232	−0.03333
15	1	2	−3.62527	−0.08294	−0.03473
17	1	3	−3.68885	−0.08395	−0.0353
19	1	4	−3.96116	−0.08559	−0.0371
21	2	3	−4.20193	−0.0731	−0.03209
15	0	3	−4.12765	−0.08894	−0.0361
23	3	2	−2.98852	−0.07692	−0.03137
25	4	1	−2.14352	−0.06539	−0.0268
15	0	3	−4.4181	−0.08737	−0.03551
17	0	4	−3.78845	−0.08891	−0.03611
17	1	5	−4.06963	−0.08798	−0.0371
19	0	5	−4.07541	−0.09274	−0.04239
21	1	3	−4.69453	−0.07144	−0.02849
15	1	3	−4.12338	−0.06524	−0.02716
23	1	3	1.700023	−0.05866	−0.0249
17	1	0	−4.06345	−0.05951	−0.12278
19	1	3	−2.55971	−0.06875	−0.13092
21	2	2	−3.11461	−0.05554	−0.11523

validation, activity of each molecule in the test set was predicted using the model generated from the training set. The *pred_r*^{2} value is calculated as follows (Eq. (2)):

$$pred\_r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{mean})^2} \quad (2)$$

where $y_i, \hat{y}_i$ are the actual and predicted activity of the *i*th molecule in the test set, respectively, and y_{mean} is the average activity of all molecules in the training set.

3. Results and discussions

The 2D- and 3D-QSAR studies of 34 Uracil-based hydroxamic acids derivatives for inhibitory histone deacetylase inhibitors through PLS and *k-nearest neighbor* methodology, respectively.

3.1. Interpretations of 2D QSAR models

pIC_{50} = 0.4820 (±0.0946) Hydrogen count −0.1235 (±0.0832) SdsCHE-index +0.0915 (±0.0046) T_2_C_1 + 1.1446 (±0.3452) SsssCHcount +0.1406 (±0.0658) $SsCH_3E$-index.

Degrees of Freedom = 20, $N_{training}$ = 27, N_{test} = 7, r^2 = 0.8498, q^2 = 0.7639, *F* test = 57.2652, r^2_se = 0.3274, q^2_se = 0.3762, *pred_r*^{2} = 0.7932, *pred_r*^{2}se = 0.3276, *Z* score $Q\hat{}2$ = 1.33392, Best Rand $Q\hat{}2$ = 0.96451.

The statistically significant penta-parametric model with GA-PLS method with coefficient of determination (r^2) = 0.8498 is capable of explaining 84.98% of variance in the observed activity values. The low standard error of r^2_se = 0.2274 demonstrates accuracy of the model. Cross-validated squared correlation coefficient of this model was 0.7639% which shows the good internal prediction power of this model. Another parameter for predictivity of test set compound is high *pred_r*^{2} = 0.7932 which is showing good external predictive power of the model.

Model 1 is obtained by the GA-PLS method which shows positive contribution of Hydrogen count, T_2_C_1, SsssCH-count, $SsCH_3E$-index and a negative correlation with SdsCHE-index. The positive coefficient associated with the $SsCH_3E$-index descriptor in the model suggests that the increased number of eCH_3 will augment the potency of the compounds. Positive contribution of this descriptor revealed the increase of histone deacetylase inhibitors of Uracil-based with the presence of CH_3 group. The higher

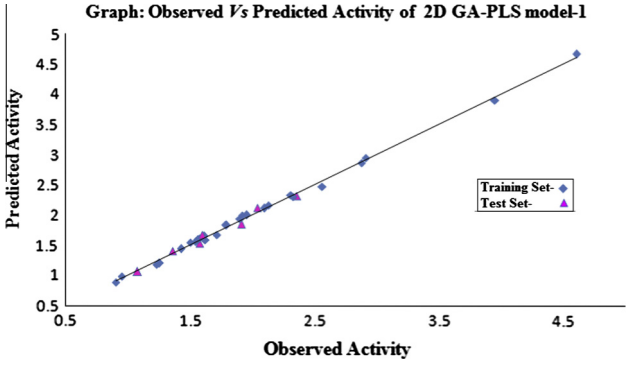


Figure 1c Plot of observed versus predicted activity by 2D QSAR GA-PLS Model-1.

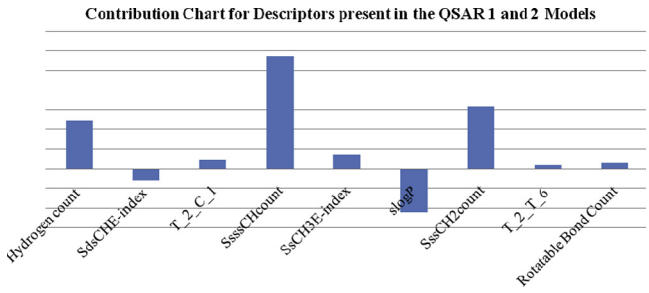


Figure 1d Contribution charts of the descriptors for the 2D Models 1 and 2.

activity of molecules 1–6, 11–21 and 26–34 compounds justifies this finding. The descriptor hydrogen count descriptor indices for number of hydrogen atoms will augment the potency of the compounds. The descriptor T_2_C_1 indicates that the presence of substituents with direct attachment of carbon on aromatic ring (i.e., —CH_3 or —C_2H_5) favorable for the activity. Positive contributions of these descriptors were clearly signifying that the presence of uracil pharmacophore was important for biological activity. The SaaCHcount [the total number of carbon atoms connected with hydrogen along with two aromatic bonds] is 28% positively contributing descriptor toward activity. SdsCHE-index is [electrotopological state index for number of —CH group connected] with one double and one single bond, i.e., aryl/alkyl atom types. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 5. The graph for observed versus predicted activity and contribution chart for Model 1 is shown in Figs. 1c and 1d respectively. The correlation matrix is shown in Table 4 which shows good correlation of selected parameters with biological activity.

pIC_{50} = −0.4486 (±0.0917) *s*log*P* + 0.6310 (±0.2815) $SssCH_2$count +0.0367 (±0.0047) T_2_T_6 + 0.0576 (±0.0037) Rotatable Bond Count −0.0036.

Degrees of Freedom = 20, $N_{training}$ = 27, N_{test} = 7, r^2 = 0.7614, q^2 = 0.6632, *F* test = 33.2158, r^2se = 0.4036, q^2se = 0.3502, *pred_r*^{2} = 0.7229, *pred_r*^{2}se = 0.5206, *Z* score $Q\hat{}2$ = 2.33343, Best Rand $Q\hat{}2$ = 1.57212.

The significant Model 2 using the GA-PLS analysis method having 0.7614 as the coefficient of determination (r^2) was considered and explains 76.14% of the variance in the observed activity values. It shows an internal predictive power (q^2 = 0.6632) of 66% and a predictivity for the external test set (*pred_r*^{2} = 0.7229) of about 72%. The GA-PLS model indicated that the descriptor $SssCH_2$count, Rotatable Bond Count, alignment-independent descriptor contribute positively T_2_T_6 and, *s*log*P* contributes negatively.

GA-PLS model reveals that the descriptor is rotatable bond count positive coefficient shows that increase in rotatable bonds in compounds is detrimental for the activity. This suggests that the presence of less number of rotatable bonds in compounds would increase the activity (like in compound 2–6; 11–21 and 26–34) in at X position of uracil moiety favored the activity. Model shows that the descriptor $SssCH_2$count plays most important role (~36%) in determining activity. The positive correlation suggests that histone deacetylase inhibitors of uracil derivatives may be increased by increasing the number of such —CH_2 groups present in the molecules. The

Please cite this article in press as: Sharma, M.C., Sharma, S. Molecular modeling study of uracil-based hydroxamic acids-containing histone deacetylase inhibitors. Arabian Journal of Chemistry (2015), http://dx.doi.org/10.1016/j.arabjc.2014.12.030

Table 4 Correlation matrix between descriptors present in the best QSAR Model-1.

Parameter	H-Count	SdsCHE-index	T_2_C_1	SsssCHcount	$SsCH_3E$-index
H-Count	1.0000				
SdsCHE-index	0.5432	1.0000			
T_2_C_1	0.3291	0.5492	1.0000		
SsssCHcount	0.3948	0.5874	0.6574	1.0000	
$SsCH_3E$-index	0.2873	0.4392	0.6148	0.7728	1.0000

Table 5 Comparative observed and predicted activities of uracil-derivatives.

Com.	pIC_{50}	2D Model-1		2D Model-3		3D Model-5		3D Model-6	
		Pred.	Res.	Pred.	Res.	Pred.	Res.	Pred.	Res.
1	2.914	2.949	−0.035	2.894	0.02	2.923	−0.009	2.902	0.012
2	1.431	1.443	−0.012	1.417	0.014	1.471	−0.04	1.402	0.029
3	0.903	0.886	0.017	0.921	−0.018	0.883	0.02	0.932	−0.029
4	1.079	1.063	0.016	1.115	−0.036	1.014	0.065	1.052	0.027
5	1.579	1.529	0.05	1.5411	0.0379	1.613	−0.034	1.541	0.038
6	1.623	1.589	0.034	1.594	0.029	1.685	−0.062	1.659	−0.036
7	2.564	2.472	0.092	2.593	−0.029	2.525	0.039	2.607	−0.043
8	1.230	1.186	0.044	1.281	−0.051	1.112	0.118	1.159	0.071
9	1.361	1.409	−0.048	1.333	0.028	1.451	−0.09	1.345	0.016
10	3.954	3.904	0.05	3.890	0.064	3.971	−0.017	3.915	0.039
11	1.579	1.614	−0.035	1.534	0.045	1.583	−0.004	1.533	0.046
12	2.359	2.316	0.043	2.394	−0.035	2.253	0.106	2.296	0.063
13	2.096	2.113	−0.017	2.137	−0.041	2.122	−0.026	2.040	0.056
14	1.255	1.204	0.051	1.261	−0.006	1.230	0.025	1.274	−0.019
15	0.954	0.990	−0.036	0.982	−0.028	0.992	−0.038	0.973	−0.019
16	1.505	1.543	−0.038	1.541	−0.036	1.432	0.073	1.461	0.044
17	1.913	1.850	0.063	1.949	−0.036	1.922	−0.009	1.874	0.039
18	1.792	1.833	−0.041	1.764	0.028	1.742	0.05	1.727	0.065
19	1.716	1.668	0.048	1.746	−0.03	1.762	−0.046	1.737	−0.021
20	1.568	1.601	−0.033	1.505	0.063	1.545	0.023	1.554	0.014
21	1.785	1.842	−0.057	1.706	0.079	1.716	0.069	1.791	−0.006
22	2.880	2.868	0.012	2.831	0.049	2.852	0.028	2.863	0.017
23	1.612	1.653	−0.041	1.663	−0.051	1.575	0.037	1.737	−0.125
24	1.903	1.936	−0.033	1.944	−0.041	1.947	−0.044	1.000	0.903
25	4.612	4.665	−0.053	4.591	0.021	4.678	−0.066	4.635	−0.023
26	1.568	1.603	−0.035	1.447	0.121	1.517	0.051	1.547	0.021
27	2.311	2.334	−0.023	2.292	0.019	2.387	−0.076	2.245	0.066
28	1.602	1.668	−0.066	1.575	0.027	1.589	0.013	1.635	−0.033
29	1.544	1.559	−0.015	1.498	0.046	1.577	−0.033	1.535	0.009
30	1.919	1.996	−0.077	1.883	0.036	1.869	0.05	1.936	−0.017
31	1.954	2.000	−0.046	1.989	−0.035	1.916	0.038	1.978	−0.024
32	2.328	2.302	0.026	2.292	0.036	2.366	−0.038	2.307	0.021
33	2.041	2.118	−0.077	2.019	0.022	2.113	−0.072	2.154	−0.113
34	2.130	2.167	−0.037	2.086	0.044	2.174	−0.044	2.091	0.039

Res. = Obs. pIC_{50} – Pred. pIC_{50}.

descriptor $s\log P$ (~24%) which is directly proportional to the activity and shows the role of thermodynamic property in determining activity. The descriptor T_2_T_6 (i.e., pair of any double bonded atom with any atom separated by six bonds) plays most important role in activity, which mainly indicates the relationship with reference to variation in different substitution patterns (mono, di, tri) on the phenyl ring.

$pIC_{50} = -0.8928$ (±0.2301) Carbons Count + 0.4796 (±0.1410) $SsCH_3E$-index + 0.8290 (±0.2880) SsssCHE + 0.1208 (±0.0534) SaasCE-index

Degrees of Freedom = 23, $N_{training} = 27$, $N_{test} = 7$, $r^2 = 0.8524$, $q^2 = 0.7750$, F test = 42.6705, r^2se = 0.3340, q^2se = 0.3011, $pred\_r^2 = 0.7447$, $pred_r^2$se = 0.4188, Z score $Q\hat{}2 = 2.42160$, Best Rand $Q\hat{}2 = 1.82987$.

Model-3 shows good correlation between biological activity and parameters Carbons Count, $SsCH_3E$-index, SsssCHE and SaasCE-index as the correlation coefficient $r^2 = 0.8524$ and the model explains about 85% variance in activity by uracil derivatives. The model shows an internal predictive ($q^2 = 0.7750$) of 77% and a predictivity for the external test ($pred\_r^2 = 0.7447$) of 74%.

SA-PLS model reveals that the SaasCE-index [electrotopological state index for number of carbon atoms connected with one single bond along with two aromatic bonds] positive

Please cite this article in press as: Sharma, M.C., Sharma, S. Molecular modeling study of uracil-based hydroxamic acids-containing histone deacetylase inhibitors. Arabian Journal of Chemistry (2015), http://dx.doi.org/10.1016/j.arabjc.2014.12.030

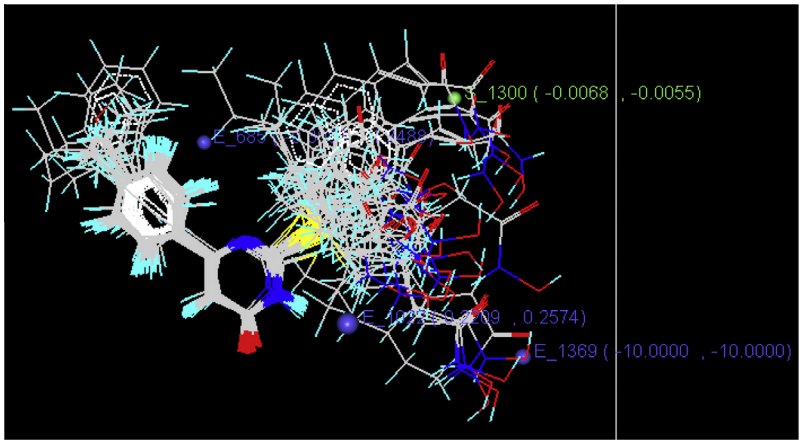


Figure 1e Contribution plot for steric and electrostatic interactions 3D QSAR GA-PLS model.

correlation (~16%) of the descriptor in the model indicates that the carbon atoms connected with aromatic rings and single bonds positively influence activity shown by substituted uracil derivatives. The SsssCHE (~19.40%) is a topological index, and signifies total number of ═CH groups connected with three single bonds. The descriptor suggests that activity can be achieved by increasing the saturated rings and saturated aliphatic chains. The positive coefficient of $SsCH_3E$-index (~26.66%) showed that increase methyl group is detrimental for the inhibitory activity against Histone deacetylase (like in compounds 1–6; 11–21 and 26–34). The descriptor Carbon Count signifies the number of carbon atoms in a compound. This provides carbon atoms in the molecule with a particular arrangement that are responsible for the activity. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 5.

pIC_{50} = 1.5252 (±0.2431) SssOE-index −0.0408 (±0.0130) Quadrupole2 + 0.3485 (±0.1179) $SssCH_2E$-index.

Degrees of Freedom = 20, $N_{training}$ = 27, N_{test} = 7, r^2 = 0.7325, q^2 = 0.6879, F test = 14.7508, r^2se = 0.3316, q^2se = 0.4572, $pred\_r^2$ = 0.7008, $pred_r^2$se = 0.2211, Z score $Q\hat{}2$ = 1.58084, Best Rand $Q\hat{}2$ = 1.18505.

Model 4 generated using SA-PLS method with 0.7325, as the coefficient of determination (r^2) was considered using the same molecules in the test and training sets. The model can explain 73% of the variance in the observed activity values. The model shows an internal predictive power (q^2 = 0.6879) of 69% and predictivity for the external test set ($pred\_r^2$ = 0.7008) of about 65%. Model 4 also shows a positive correlation with SssOE-index, $SssCH_2E$-index, and a negative correlation with Quadrupole2.

The descriptor SssOE-index which is electrotopological state indices for number of oxygen atom connected with two single bonds showed positive contribution with contribution of ~48.15%. This diminishing influence is augmented by the presence of O-Me groups at the R position. The other descriptor $SssCH_2E$-index, which signifies estate contributions defining electrotopological state indices for the number of CH_2 groups attached to two single bonds, also showed a positive contribution (~25.92%). The descriptor Quadrupole2 which carries a negative sign in this model meaning that the decreased Quadrupole of the molecule decreases the activity.

3.2. Interpretations of 3D QSAR and pharmacophore model

In present investigation, two widely used techniques, viz., genetic algorithm (GA) and simulated annealing (SA) have been applied for descriptor optimization.

The q^2, $pred\_r^2$, Vn and k value of kNN-MFA with GA and SA were (0.6790, 0.7339, 4/4) and (0.6538, 0.6043, 3/4) although there are no common descriptors among these two methods. Genetic algorithm kNN-MFA method have better q^2 (0.6790) and $pred\_r^2$ (0.7339) than other method, model validation correctly predicts activity 68% and 73.3% for the training and test set respectively.

$$k\ nearest\ neighbor = 4,\ N_{training} = 27,\ N_{test} = 7,\ q^2 = 0.6790,\ q^2\text{se} = 0.4382,\ pred\_r^2 = 0.7339,\ pred\_r^2\text{se} = 0.3287.$$

From 3D-QSAR model, it is observed that electrostatic descriptors like E_685 (~20%) with negative coefficient are near the X position of the Uracil ring (Fig. 1e). This indicates that negative electronic potential is required to increase activity and more electronegative substituents group is preferred in that X position. Electrostatic descriptors like E_1022 with positive coefficients are at the X ring of Uracil structure indicating that electropositive groups are favorable on this site and

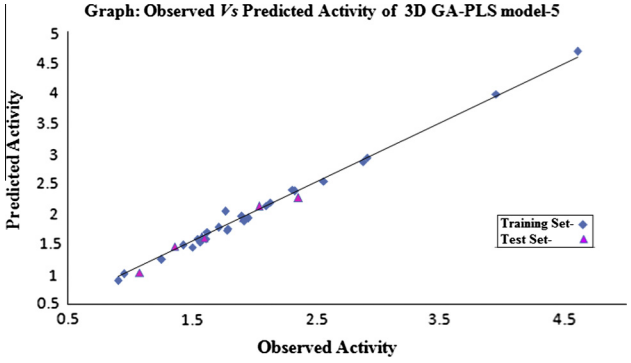


Figure 1f Plot of observed versus predicted activity by best 3D QSAR model.

Please cite this article in press as: Sharma, M.C., Sharma, S. Molecular modeling study of uracil-based hydroxamic acids-containing histone deacetylase inhibitors. Arabian Journal of Chemistry (2015), http://dx.doi.org/10.1016/j.arabjc.2014.12.030

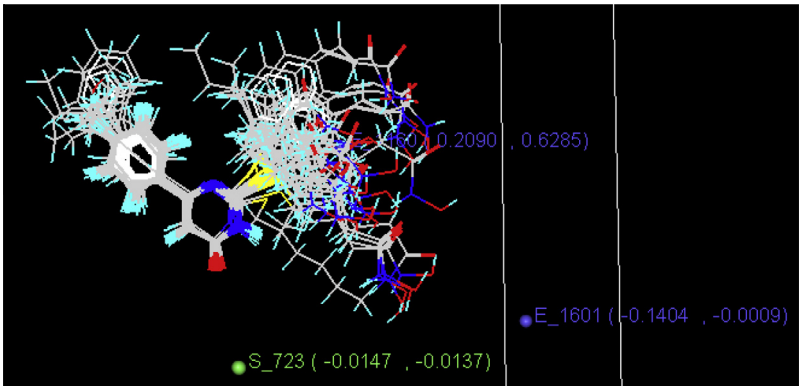


Figure 1g Contribution plot for steric and electrostatic interactions SA-PLS model.

the presence of electropositive groups would increase the histone deacetylase inhibitors activity of these compounds. Most of the compounds (compounds 1–6, 10–21, 25–34, etc.) with higher activity having electropositive substitution at the X position of uracil ring strongly support the above statement. This indicates that that group that imparting positive electrostatic potential is favorable for activity so less electronegative group is preferred in that region. The plots of Predicted versus Observed values of pIC_{50} are shown in Fig. 1f). The presence of steric descriptors S_1300 with negative coefficients is also near from the X position of the ring which indicates that less bulky groups are unfavorable on this site and presence of less bulky groups decreases the histone deacetylase inhibitors activity of Uracil compounds.

$$k\ nearest\ neighbor = 4,\ N_{training} = 27,\ N_{test} = 7,\ q^2 = 0.6838,\ q^2\text{se} = 0.2165,\ pred\_r^2 = 0.6343,\ pred\_r^2\text{se} = 0.5063$$

Model 6 (SA-kNN-MFA) is used for internal predictivity, the value of LOO cross-validation squared correlation coefficient suggested goodness of the prediction. It is observed that steric, electrostatic descriptors like S_723 and *E*_1601 with negative coefficients are from at the X and R uracil moiety (Fig. 1g). The observed and the predicted biological activity obtained using *k-nearest neighbor* model for the training and the test set compounds are shown in Table 5.

We were generated different pharmacophore patterns based on a set thirty-four aligned molecules. Selected pharmacophore shows four chemical features which were present in all 34 molecules. The information shows that the four features used were one AroC feature (Aromatic), one AlaC (aliphatic) and two HAc (Hydrogen bond acceptor) features. The average RMSD of the pharmacophore alignment of each two molecules is 0.0754 Å (Fig. 1h).

4. Conclusion

The genetic algorithm and simulated annealing are applied to the optimization and selection of suitable descriptors for the development of QSAR models for uracil based derivatives. The objective of the present multiple QSAR investigations was to develop Pharmacophore, 3D QSAR models based on similarity indices and 2D QSAR models based on classical descriptors. The QSAR studies results obtained from the study provide significant statistical parameters and various valida-

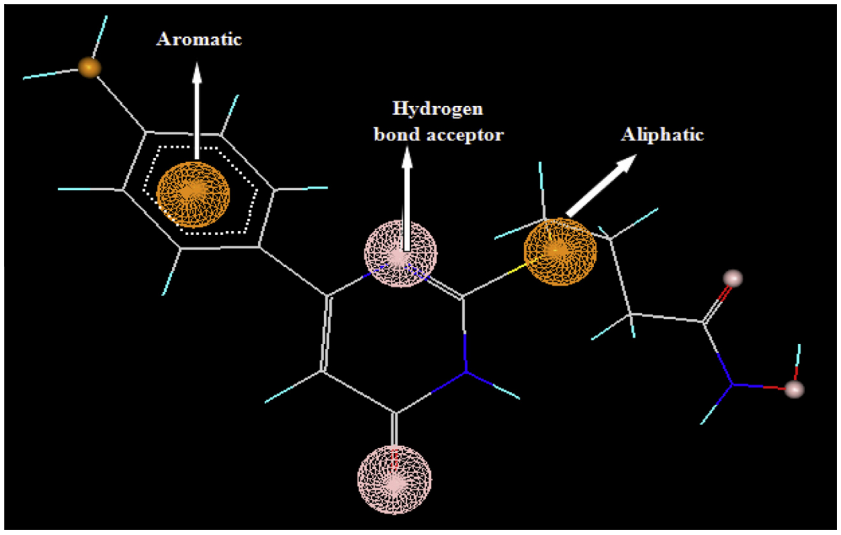


Figure 1h Best molecules pharmacophore sites.

Please cite this article in press as: Sharma, M.C., Sharma, S. Molecular modeling study of uracil-based hydroxamic acids-containing histone deacetylase inhibitors. Arabian Journal of Chemistry (2015), http://dx.doi.org/10.1016/j.arabjc.2014.12.030

tion studies confirm that the training set models 1–6 are statistically reliable and robust. The 2D-QSAR model reported herein provides some interesting insight into understanding the electrotopological descriptors, molecular connectivity descriptors and hydrophobicity descriptor are contributed significantly for the stability of the models. Presence of methoxy groups at R position of uracil moiety favored the activity. This consideration led us to further develop partial least squares analysis models with classical descriptors.

Acknowledgments

The author wishes to express gratitude to Dr. Amit Bedi V-life Science Technologies Pvt. Ltd. for providing the software for the study, and Head, School of Pharmacy, Devi Ahilya Vishwavidyalaya Indore (M.P.)-India for providing environment and facilities to carry out the work.

References

Ajmani, S., Jadhav, K., Kulkarni, S.A., 2006. Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. J. Chem. Inf. Model. 46, 24–31.

Baumann, K., 2002. An alignment-independent versatile structure descriptor for QSAR and QSPR based on the distribution of molecular features. J. Chem. Inf. Comput. Sci. 42, 26–35.

Bouchain, G., Frechette, S., Woo, S.H., Khalil, E.A., Leit, S., Fournel, M., Yan, P.T., Trachy-Bourget, M.C., Beaulieu, C., Li, Z., Besterman, J., Delorme, D., 2001. Design and synthesis of a novel class of histone deacetylase inhibitors. Bioorg. Med. Chem. Lett. 11, 2847–2850.

Clark, M., Cramer III, R.D., Van, O.N., 1989. Validation of the general purpose Tripose 5.2 Force Field. J. Comput. Chem. 10, 982–1012.

Cramer III, R.D., Bunce, J.D., Patterson, D.E., 1988. Cross validation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. Quant. Struct. Act. Rel. 7, 18–25.

Curtin, M., Glaser, K., 2003. Histone deacetylase inhibitors: the Abbott experience. Curr. Med. Chem. 10, 2373–2392.

Garea, V.A., Esteller, M., 2004. Histone deacetylase inhibitors; understanding a new wave of anticancer agents. Int. J. Cancer 112, 171–178.

Gasteiger, J., Marsili, M., 1980. Iterative partial equalization of orbital electronegativity-a rapid access to atomic charges. Tetrahedron 36, 3219–3228.

Golbraikh, A., Tropsha, A., 2002. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. J. Comput.-Aided Mol. Des. 16, 357–369.

Gregoretti, I.V., Lee, Y.M., Goodson, H.V., 2004. Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. J. Mol. Biol. 338, 17–31.

Grozinger, C.M., Schreiber, S.L., 2002. Deacetylase enzymes: biological functions and the use of small-molecule inhibitors. Chem. Biol. 9, 3–16.

Halgren, T.A., 1996. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. J. Comput. Chem. 17, 520–552.

Han, J.W., Ahn, S.H., Park, S.H., Wang, S.Y., Bae, G.U., Seo, D.W., Known, H.K., Hong, S., Lee, Y.W., Lee, H.W., 2000. Apicidin, a histone deacetylase inhibitor, inhibits proliferation of tumor cells via induction of p21WAF1/Cip1 and gelsolin. Cancer Res. 60 (21), 6068–6074.

Inche, A.G., La Thangue, N.B., 2006. Keynote review: chromatin control and cancer-drug discovery: realizing the promise. Drug Discov. Today 11, 97–109.

Insinga, A., Monestiroli, S., Ronzoni, S., Gelmetti, V., Marchesi, F., Viale, A., Altucci, L., Nervi, C., Minucci, S., Pelicci, P.G., 2005. Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. Nat. Med. 11, 71–76.

Johnstone, R.W., 2002. Histone-deacetylase inhibitors: novel drugs for the treatment of cancer. Nat. Rev. Drug Discov. 1, 287–299.

Kijima, M., Yoshida, M., Suguta, K., Horinouchi, S., Beppu, T., 1993. Trapoxin, an antitumor cyclic tetrapeptide, is an irreversible inhibitor of mammalian histone deacetylase. J. Biol. Chem. 268 (30), 22429.

Mai, A., Massa, S., Ragno, R., Esposito, M., Sbardella, G., Nocca, G., Scatena, R., Jesacher, F., Loidl, P., Brosch, G., 2002. Binding mode analysis of 3-(4-benzoyl-1-methyl-1H-2- pyrrolyl)-N-hydroxy-2-propenamide: a new synthetic histone deacetylase inhibitor inducing histone hyperacetylation, growth inhibition, and terminal cell differentiation. J. Med. Chem. 45, 1778–1784.

Mai, A., Massa, S., Rotili, D., Simeoni, S., Ragno, R., Botta, G., Nebbioso, A., Miceli, M., Altucci, L., Brosch, G., 2006. Synthesis and biological properties of novel, uracil-containing histone deacetylase inhibitors. J. Med. Chem. 49, 6046–6056.

Marks, P.A., Rifkind, R.A., Richon, V.M., Breslow, R., Miller, T., Kelly, W.K., 2001. Histone deacetylases and cancer: causes and therapies. Nat. Rev. Cancer 1, 194–202.

Monneret, C., 2005. Histone deacetylase inhibitors. Eur. J. Med. Chem. 40, 1–13.

Nebbioso, A., Clarke, N., Voltz, E., Germain, E., Ambrosiino, C., Bontempo, P., Alvarez, R., Schiavone, E.M., Ferrara, F., Bresciani, F., Weisz, A., Lera, A.R.D., Gronemeyer, H., Altucci, L., 2005. Tumor-selective action of HDAC inhibitors involves TRAIL induction in acute myeloid leukemia cells. Nat. Med. 11, 77–84.

Pandolfi, P.P., 2001. Transcription therapy for cancer. Oncogene 20, 3116–3127.

Pazin, M.J., Kadonaga, J.T., 1997. What's up and down with histone deacetylation and transcription? Cell 89, 325.

Richon, V.M., Emiliani, S., Verdin, E., Webb, Y., Breslow, R., Rifkind, R.A., Marks, P.A., 1998. A class of hybrid polar inducers of transformed cell differentiation inhibits histone deacetylases. Proc. Natl. Acad. Sci. U.S.A. 95 (6), 3003–3007.

Somech, R., Izraeli, S., Simon, J.A., 2004. Histone deacetylase inhibitors – a new tool to treat cancer. Cancer Treat. Rev. 30, 461–472.

Sotriffer, C.A., Winger, R.H., Liedl, K.R., Rode Bernd, M., Varga, J.M., 1996. Comparative docking studies on ligand binding to the multispecific antibodies IgE-La2 and IgE-Lb4. J. Comput.-Aided Mol. Des. 10 (4), 305.

Verdin, E., Dequiedt, F., Kasler, H.G., 2003. Class II histone deacetylases: versatile regulators. Trends Genet. 19, 286–293.

VLife MDS 3.5, 2008. Molecular Design Suite. Vlife Sciences Technologies Pvt. Ltd., Pune, India.

Weinmann, H., Ottow, E., 2004. Recent advances in medicinal chemistry of histone deacetylase inhibitors. Ann. Rep. Med. Chem. 39, 185–196.

Yoshida, M., Kijima, M., Akita, M., Beppu, T., 1990. Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. J. Biol. Chem. 265 (28), 17174–17179.

Please cite this article in press as: Sharma, M.C., Sharma, S. Molecular modeling study of uracil-based hydroxamic acids-containing histone deacetylase inhibitors. Arabian Journal of Chemistry (2015), http://dx.doi.org/10.1016/j.arabjc.2014.12.030